

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-344905
(P2001-344905A)

(43) 公開日 平成13年12月14日 (2001. 12. 14)

(51) Int.Cl. ⁷	識別記号	F I	テームト [*] (参考)
G 1 1 B 20/10	3 2 1	G 1 1 B 20/10	3 2 1 Z 5 D 0 4 4
G 1 0 L 19/02		G 1 0 L 7/04	G 5 D 0 4 5
19/00		9/18	M

審査請求 未請求 請求項の数 5 O L (全 15 頁)

(21) 出願番号 特願2000-157042(P2000-157042)

(22) 出願日 平成12年5月26日 (2000. 5. 26)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 安孫子 幸弘

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 加藤 秀雄

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 100074099

弁理士 大曾 義之 (外1名)

最終頁に続く

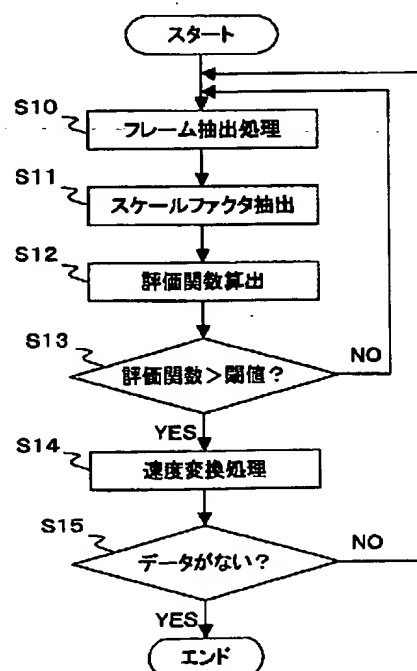
(54) 【発明の名称】 データ再生装置、その方法及び記録媒体

(57) 【要約】

【課題】 MPEGデータを簡単な構成で話速変換可能な再生装置及び方法を提供する。

【解決手段】 MPEGオーディオデータをデコードしないまま、そのデータ単位であるフレームを抽出する。そして、フレームに含まれるスケールファクタを抽出し、スケールファクタに基づいて評価関数を算出する。そして、評価関数の値が所定の閾値より大きければ、そのフレームを速度変換処理する。評価関数の値が所定の閾値より小さければ、無音区間のフレームであるとして無視する。速度変換処理は、フレームを所定の法則に基づいて間引いたり、同じフレームを所定回数繰り返すことにより行う。

本発明の話速変換処理の概略フローチャート



【特許請求の範囲】

【請求項 1】オーディオデータを含む圧縮されたマルチメディアデータを再生する装置であって、
該オーディオデータの単位データであるフレームを抽出する抽出手段と、

該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行う変換手段と、
該変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備え、
た、ことを特徴とするデータ再生装置。

【請求項 2】オーディオデータを含むマルチメディアデータを再生する装置であって、

該オーディオデータの単位データであるフレームを抽出する抽出手段と、

該オーディオデータの再生速度を設定する設定手段と、
該設定手段によって設定された該再生速度に基づいて、
該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行う速度変換手段と、
該速度変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備え、

圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とするデータ再生装置。

【請求項 3】前記フレームに含まれるスケールファクタを抽出するスケールファクタ抽出手段と、

該スケールファクタに対して演算を行う演算手段と、
該演算手段の演算結果と所定の閾値とを比較し、該演算結果が該閾値よりも小さい場合には、対応する該フレームを前記再生手段に送らないように制御する制御手段と、
を更に備えることを特徴とする請求項 2 に記載のデータ再生装置。

【請求項 4】オーディオデータを含むマルチメディアデータを再生する方法であって、

(a) 該オーディオデータの単位データであるフレームを抽出するステップと、

(b) 該オーディオデータの再生速度を設定するステップと、

(c) 該ステップ (b) において設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行うステップと、

(d) 該ステップ (c) の処理後受け取った該オーディオデータの該フレームをデコードし、音声を再生するステップとを備え、

圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とするデータ再生方法。

【請求項 5】情報装置に、オーディオデータを含むマルチメディアデータを再生する方法を実行させるプログラ

ムを記録した記録媒体であって、

(a) 該オーディオデータの単位データであるフレームを抽出するステップと、

(b) 該オーディオデータの再生速度を設定するステップと、

(c) 該ステップ (b) において設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行うステップと、

10 (d) 該ステップ (c) の処理後受け取った該オーディオデータの該フレームをデコードし、音声を再生するステップとを備え、

圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とする方法を実現するプログラムを記録した、情報装置読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データ再生装置及び再生方法に関する。

【0002】

【従来の技術】近年のデジタル音声録音技術の発展により、従来のテープレコーダに代わり、MD レコーダを使用して、MD に音声を録音することが一般的となっている。また、更に、現在では、DVD などを従来のビデオテープの代わりに使用し、映画などを一般に配布することが行われ始めている。このようなデジタル音声録音技術、及び画像録画技術には、様々な技術が使用されるが、そのような技術の中でも MPEG は最も一般的な技術の一つである。

【0003】図 15 及び図 16 は、MPEG オーディオデータのフォーマットを示す図である。MPEG オーディオデータは図 15 に示されるように、AAU (Audio Access Unit あるいは Audio Frame) と呼ばれる連続したフレームで構成される。フレームは、更にヘッダ、エラーチェック、オーディオデータ、アンシラリデータ (付加情報) で構成される階層構造を持つ。ここで、オーディオデータは、圧縮処理されたデータである。

【0004】ヘッダは、syncword、レイヤ、ビットレートに関する情報、サンプリング周波数に関する情報、パディングビットなどのデータで構成される。この構造は圧縮性能の異なるレイヤ I、レイヤ II、レイヤ III で共通している。

【0005】フレーム中のオーディオデータは図 16 のように構成されている。図 16 に示されるように、オーディオデータは、レイヤ I、II、III に関わらず、必ずスケールファクタを含んでいる。このスケールファクタは波形の再生時の倍率を表したデータである。すなわち、レイヤ I、II のサンプリングデータあるいは、レイヤ III のハフマンコードビットが表す音声デ

ータは、スケールファクタで正規化されており、実際の音声データは、サンプリングデータや、ハフマンコードビットを伸長したデータにスケールファクタを乗算して得られる。スケールファクタは時間軸に沿って更に32区間（サブバンド）に分かれて圧縮されており、スケールファクタもモノラルで最大32個割り当てられる。

【0006】なお、MPEGオーディオの詳細については、国際規格であるISO/IEC11172-2を参照されたい。図17は、従来のMPEGオーディオ再生装置の基本的構成を示す図である。

【0007】MPEGオーディオデータがMPEGオーディオ入力部10に入力されると、上記国際規格に記載されている処理を実現するMPEGオーディオ復号化部11において復号され、スピーカなどからなるオーディオ出力部12から音声として出力される。

【0008】

【発明が解決しようとする課題】デジタルで録音された音声を再生する場合、再生速度を変更することは頻繁に行われ、特に、話速変換は内容理解や内容圧縮に有効である。ところが、MPEGオーディオデータから直接話速変換する場合、従来においては、一旦デコードしてから話速変換を行っていた。

【0009】MPEGオーディオにおける圧縮は、データを数十分の1まで低減させることができる。したがって、MPEGオーディオデータをデコードしてから話速変換処理を行おうとすると、圧縮データを伸長した後の膨大なデータを扱わなければならない。従って、話速変換に必要な回路の数や規模が大きくなってしまう。

【0010】MPEGオーディオデータを一旦デコードしてから話速変換するものである公知例として特開平9-73299号公報がある。本発明の課題は、マルチメディアデータを簡単な構成で話速変換可能な再生装置及び方法を提供することである。

【0011】

【課題を解決するための手段】本発明の第1のデータ再生装置は、オーディオデータを含む圧縮されたマルチメディアデータを再生する装置であって、該オーディオデータの単位データであるフレームを抽出する抽出手段と、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行う変換手段と、該変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備えたことを特徴とする。

【0012】本発明の第2のデータ再生装置は、オーディオデータを含むマルチメディアデータを再生する装置であって、該オーディオデータの単位データであるフレームを抽出する抽出手段と、該オーディオデータの再生速度を設定する設定手段と、該設定手段によって設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力

処理を行う速度変換手段と、該速度変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備え、圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とする。

【0013】本発明のデータ再生方法は、オーディオデータを含むマルチメディアデータを再生する方法であって、(a)該オーディオデータの単位データであるフレームを抽出するステップと、(b)該オーディオデータの再生速度を設定するステップと、(c)該ステップ(b)において設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行うステップと、(d)該ステップ(c)の処理後受け取った該オーディオデータの該フレームをデコードし、音声を再生するステップとを備え、圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とする。

【0014】本発明によれば、圧縮されたオーディオデータを、デコードすることなく、圧縮されたままで、話速変換処理を行うことが可能となるので、データ再生装置に必要とされる回路規模が縮小され、簡単な構成で話速変換を伴う再生処理を行うことが出来る。

【0015】

【発明の実施の形態】本発明の実施形態は、MPEGオーディオからオーディオフレームと呼ばれるフレームを抽出し、フレームを所定の法則に従って間引くことによって変換速度を速める、あるいはフレームを所定の法則に従って内挿することによって変換速度を遅くする。また、抽出したフレームから得られるスケールファクタを用いて評価関数を算出し、所定の法則に従って当該フレームの間引くことによって無音区間を圧縮する。更に、接続される前後のフレームでスケールファクタを変更することで、接続部での聴覚上の違和感（ノイズなど）を低減する。また再生装置においては、データ入力部とMPEGデータ判別部と上記方法によって話速変換変換を行う話速変換部とMPEGオーディオ部とオーディオ出力部を備える。

【0016】本発明の実施形態で行うフレーム抽出について、図16、17のMPEGオーディオデータの構成を参照して説明する。フレームの抽出は、フレームの先頭にあるsyncwordを検出することによって行う。つまり、フレームnのsyncwordの先頭からフレームn+1のsyncwordの直前までのビット列を読み込む。

【0017】あるいは、別なフレームの抽出方法として、syncwordを含む32ビットのビット列からなるオーディオフレームヘッダから、ビットレート、サンプリング周波数、パディングビットを抽出し、次式から1フレームのデータ長を求め、syncwordからデータ長までのビット列を読み込んで良い。

{フレームサイズ×ビットレート [bit/sec] ÷ 8 ÷ サ

ンプリング周波数 [Hz] } + パディングビット [byte]

話速変換では、再生速度を変化させたときに聞き手に違和感を感じさせないことが重要であるため、次にステップで処理を行う。

- ・基本周期の抽出
- ・基本周期の間引き、繰り返し
- ・無音部分の圧縮

音声の周期性を持つ波形の周期を基本周期と呼び、日本人男性で100~150Hz、女性で250~300Hzである。話速を速くする場合には周期性のある波形を抽出し間引く処理、一方、遅くする場合には逆に繰り返す処理を行う。

【0018】MPEGオーディオに対して従来の話速変換方法を適用する場合には、次の問題がある。

- ・PCM形式への復元が必要
- ・リアルタイム処理には専用のハードウェアが必要

音声処理においては、その処理単位として約10~30ミリ秒程度が一般的である。MPEGオーディオでは1オーディオフレームは約20ミリ秒（レイヤII、44.1kHz、1152サンプルの場合）である。

【0019】この基本周期をオーディオフレームで代用することで復元なしに話速変換の効果を得る。一方、従来無音区間を検出するためには、音圧の強弱を評価する必要があった。厳密にはデコードなしに無音区間を正確に検出することはできない。しかし、オーディオデータに含まれるスケールファクタは波形の再生の倍率を表したデータであるので、音圧と近い性質を持ち、本実施形態ではスケールファクタを利用する。

【0020】図1は、同じ音声データをMPEGオーディオレイヤIIで圧縮したデータのスケールファクタと圧縮しないデータの音圧について比較した図である。グラフの縦軸は1フレーム（MPEGオーディオレイヤII相当：1152サンプル）についてのスケールファクタの平均、あるいは音圧の区間平均を表し、横軸は時間を表す。スケールファクタと音圧が非常に近い形状を示しており、この図においては相関係数は約80%程度であり高い相関を示している。エンコードの性能にも依存するが、スケールファクタは、音圧と近い性質を持っていることが示されている。

【0021】従って、本実施形態においては、スケールファクタから評価関数を演算することによって無音区間の検出をする。評価関数の一例としては1フレーム内のスケールファクタの平均値とすることが考えられる。あるいは、数フレームにわたって評価関数を設定してもよいし、サブバンド毎のスケールファクタから評価関数を設定してもよく、更にこれらを組み合わせたものでもよい。

【0022】ところで、単にフレーム単位に間引きを行い接続する場合、フレームとフレームの接続点で聴感上

違和感を感じる場合がある。これは、間引きを行ったことにより、音圧の変換が不連続に大きくなったり、小さくなることによって生じる。そこで、本実施形態では、フレームとフレームの接続点前後のフレームの一部のスケールファクタを変更することで、この違和感を低減する。

【0023】例えば、接続点直前のスケールファクタが0に近く、かつ接続点直後のスケールファクタが最大値に近い場合、接続部では本来含まれない高周波成分が加わってしまい。これがノイズとして聴感上の違和感となって現れることになる。この場合については、接続点前後のスケールファクタを中間値に変更することで低減される。

【0024】本発明の実施形態においては、MPEGオーディオ規格で定義されているオーディオフレームと呼ばれるフレームを単位としてMPEGデータをデコードせずに話速変換するので、MPEGデータをデコードする必要が無く、回路規模を縮小し、簡単な構成で話速変換を実現することが出来る。また、スケールファクタを用いることで、デコードによって音圧を得ることなしに無音区間を判定して、無音区間を削除し、有音区間を割り当てることにより話速変換を行う。更に、スケールファクタを適切に変換することで、フレームの接続点前後の聴感上の違和感を低減することができる。

【0025】図2は、本発明の話速変換処理の概略フローチャートである。まず、ステップS10において、フレームの抽出を行う。フレームの抽出は、フレームの先頭にあるsyncwordを検出することで行う。つまり、フレームnのsyncwordの先頭からフレームn+1のsyncwordの直前までのビット列を読み込む用にする。あるいは、syncwordを含む32ビットのビット列からなるオーディオフレームヘッダから、ビットレート、サンプリング周波数、パディングビットを抽出し、前述の式から1フレームのデータ長を求め、syncwordからデータ長までのビット列を読み込んでよい。また、フレームの抽出は、MPEGオーディオの復号化では必須の処理であるので、MPEGオーディオ復号化で用いられるフレームの抽出機能をそのまま利用することで実現してもよい。そして、フレーム抽出が成功すると、次に、スケールファクタの抽出を行う。図17に示されるように、スケールファクタは、MPEGオーディオのオーディオデータの先頭から、各レイヤにおいて、決められたビット位置に存在するので、syncwordからのビット数を係数する事によって、スケールファクタを抽出することができる。あるいは、スケールファクタの抽出も、フレームの抽出と同様にMPEGオーディオの復号化では必須の処理であるので、既存のMPEGオーディオ復号化処理によって抽出されたスケールファクタを使用してもよい。

【0026】次に、ステップS12において、スケールファクタから評価関数算出処理を行う。評価関数はスケ

7

ールファクタから算出する。評価関数の簡単な一例としては1フレーム内のスケールファクタの平均値とすることが考えられる。あるいは、数フレームにわたって評価関数を設定しても良いし、サブバンド毎のスケールファクタから評価関数を設定してもよく、さらにこれらを組み合わせたものでも良い。

【0027】次に、評価関数の演算値と予め決定された閾値との比較を行う。閾値より評価関数の方が大きい場合には、有音区間のフレームであるとしてステップS14に進む。また、評価関数の値が閾値以下の場合には、無音区間のフレームであるとしてフレームを無視し、ステップS10に戻る。ここで、閾値は固定でも、可変でもよい。

【0028】ステップS14においては、速度変換処理を行う。速度変換処理は、オリジナルのMP EGデータの再生速度を1としたとき、所望の再生速度が1より大きい場合、ある間隔でフレームを間引き、データを短縮して出力する。例えば、フレームを先頭から0、1、2、・・・とした場合、2倍速である場合には、0、2、4、・・・等のように、フレームを間引いて復号化し、再生する。また、所望の再生速度が1未満の場合、ある間隔でフレームを繰り返して出力する。例えば、上記例を採用すると、1/2倍速の時は、フレームを0、0、1、1、2、2、・・・というようにフレームを配列し、復号化して再生する。聴者は、このようにして出力されたMP EGデータがデコードされると所望の速度で再生されたように聞こえる。

【0029】そして、ステップS14において、速度変換処理があるフレームについて終了すると、ステップS15に進んで、まだ処理すべきデータが存在するか否かを判断し、データが存在する場合には、ステップS10に戻って、次のフレームを処理し、データがない場合には、処理を終了する。

【0030】図3は、本発明の別の話速変換処理の概略フローチャートである。まず、図2の場合と同様に、ステップS20において、フレーム抽出処理を行い、ステップS21において、スケールファクタ抽出処理を行う。そして、ステップS22において、評価関数を演算し、ステップS23において、評価関数の値と閾値との比較を行う。ステップS23において、評価関数の方が閾値よりも大きいと判断された場合には、有音区間のフレームであるとして、ステップS24に進む。ステップS23において、評価関数の値が閾値以下であると判断される場合には、無音区間のフレームであるとして、ステップS20に戻って、次のフレームの処理をする。

【0031】ステップS24においては、図2で説明したような速度変換処理を行い、更に、ステップS25において、フレームの接続部におけるノイズの発生を抑えるため、スケールファクタの修正処理を行う。そして、ステップS26において、後続のデータがあるか否かを

8

判断し、データがある場合には、ステップS20に進み、データがない場合には処理を終了する。なお、スケールファクタ修正処理では、直前のフレームを保持しており、フレームの接続点前後のスケールファクタを調整し出力する。

【0032】図4は、再生速度変換処理の流れを示す詳細なフローチャートである。なお、同図においては、 n_{in} を入力フレーム数とし、 n_{out} を出力フレーム数とし、 K を再生速度とする。

【0033】まず、ステップS30において、初期化する。すなわち、 n_{in} を-1に、 n_{out} を0に設定する。次に、ステップS31において、オーディオフレームの抽出処理を行う。この処理は、前述したように、既存の技術を使用して実現することができるので、詳細には説明しない。次に、ステップS32において、オーディオフレームの抽出処理が正常に行われたか否かを判断する。ステップS32において、オーディオフレームの抽出が正常に行われなかったと判断された場合には、処理を終了する。ステップS32において、オーディオフレームの抽出が正常に行われたと判断された場合には、ステップS33に進む。

【0034】ステップS33においては、入力フレーム数である n_{in} を1だけ増加する。そして、ステップS34において、再生速度 K が1以上であるか否かが判断される。この再生速度は、一般には、再生装置を使用するユーザによって設定されるものである。ステップS34において、再生速度 K が1以上であると判断された場合には、出力フレーム数 n_{out} の再生速度 K 倍が入力フレーム数 n_{in} 以上になったか否かを判断する（ステップS35）。すなわち、入力フレームを間引いて出力される出力フレームの数の再生速度 K 倍が入力フレーム数 n_{in} 以下か否かを判断する。ステップS35の判断がNOとなった場合には、ステップS31に戻り、ステップS35の判断がYESとなった場合には、ステップS36に進む。

【0035】ステップS36においては、オーディオフレームを出力する。そして、ステップS37において、出力フレーム数 n_{out} を1だけ増加し、ステップS31に戻る。

【0036】図4の K が1以上である場合の、オーディオフレームの出力は処理を繰り返すことによって間引き処理を行うようになっている。間引き方としては、前述の2倍速の他に、3倍速の場合においては、0、3、6、・・・などを行う。1.5倍速などの場合には、整数 N 、 M に対して、 $1.5 \times N = M$ を計算し、 M 番のフレームを $N+1$ 番目に配置し、このようにして配置されたフレームの間に適当なフレームを埋め込むことによって行う。すなわち、1.5倍の場合、0、1、3、4、6、・・・とするか、0、2、3、5、6、・・・などを行う。

【0037】ステップS34において、再生速度Kが1より小さい場合には、ステップS38において、オーディオフレーム出力処理を行う。この場合は、フローに従ってオーディオフレームを出力することによって、例えば、1/2倍速の場合には、0、0、1、1、2、2、・・・のように、1/3倍速の場合には、0、0、0、1、1、1、2、2、2、・・・のようにフレームを繰り返して出力し、1より小さい再生速度を実現する。

【0038】そして、ステップS39において、出力フレーム数 n_{out} を1だけ増加し、ステップS40において、入力フレーム数 n_{in} が出力フレーム数 n_{out} の再生速度K倍以下であるか以下を判断する。ステップS40の判断がYESの場合には、ステップS41に戻る。ステップS40の判断がNOの場合には、ステップS38に戻って、同じフレームを繰り返して出力する。

【0039】以上のような処理を繰り返して、再生速度変換処理を行う。図5は、再生速度変換処理及び無音部分除去処理を含む処理の流れを示す詳細フローチャートである。

【0040】まず、ステップS45において、 n_{in} を-1に、 n_{out} を0に初期化する。次に、ステップS46において、オーディオフレームの抽出処理を行う。ステップS47において、オーディオフレームの抽出処理が正常か否かを判断する。オーディオフレームの抽出処理が異常であった場合には、処理を終了する。オーディオフレームの抽出処理が正常であった場合には、ステップS48に進んで、スケールファクタの抽出を行う。スケールファクタの抽出方法は、前述したように、既存の技術で行うことができるので、詳細な説明は省略する。そして、ステップS49に進んで、抽出したスケールファクタから評価関数F（前述したように、例えば、1フレーム分のスケールファクタの和）を算出する。そして、ステップS50において、入力フレーム数 n_{in} を1増加してステップS51に進む。ステップS51においては、 $n_{in} \geq K \cdot n_{out}$ かつ $F > Th$ （ここで、Thは閾値）であるか否かを判断する。ステップS51において、判断がNOの場合には、ステップS46に戻る。ステップS51における判断がYESの場合には、ステップS52に進んで、オーディオフレームを出力し、ステップS53において、出力フレーム数 n_{out} を1だけ増加して、ステップS46に進む。

【0041】ここで、ステップS51の $n_{in} \geq K \cdot n_{out}$ の判断式の意味は、図4で説明したものと同様である。また、 $F > Th$ も前述の概略フローによって説明したとおりである。

【0042】図6は、ノイズ低減処理の流れを示すフローチャートである。まず、ステップS60において、 n_{in} を-1、 n_{out} を0に設定して、初期化を行う。次に、ステップS61において、オーディオフレーム抽出処理を行い、ステップS62において、オーディオフレ

ーム抽出処理が成功であったか否かを判断する。オーディオフレーム抽出処理が失敗であった場合には、処理を終了する。オーディオフレーム抽出処理が成功であった場合には、ステップS63に進む。

【0043】次に、ステップS63において、スケールファクタを抽出し、ステップS64において、評価関数Fを算出する。そして、ステップS66において、入力フレーム数 n_{in} を1だけ増加し、ステップS67において、 $n_{in} \geq K \cdot n_{out}$ かつ $F > Th$ であるか否かを判断する。ステップS67の判断がNOの場合には、ステップS61に進み、ステップS67の判断がYESの場合には、ステップS68において、スケールファクタ修正処理を行う。

【0044】そして、ステップS69に進み、オーディオフレーム出力処理を行い、ステップS70において、 n_{out} を1だけ増加してステップS61に戻る。図7、及び、図8は、図6のスケールファクタ修正処理を説明する図である。

【0045】図7に示されるように、オーディオフレームを間引いたりして送出する場合、オーディオフレームの接続点において、音圧の不連続な変化が生じる。このような不連続は、音声を聞いているユーザにとっては、ノイズとして聞こえてしまい、早送りなどを行っている場合に、非常に不快な音に気が取られてしまう。

【0046】そこで、図8に示すように、オーディオフレームの境界付近で係数値が小さくなるような修正係数をスケールファクタに乗算して、音声を再生するようにする。このようにすることによって、図8の太線で示されているように、オーディオフレームの接続点付近での音圧の不連続な飛びが緩和される。従って、再生音を聞いているユーザにとっても、ノイズが小さくなり、早送りなどの場合、不快な音が気にならなくなる。

【0047】図9は、本発明の話速変換を適用したMP3オーディオデータ再生装置の一構成を示すブロック構成図である。これは図18に示す従来のMP3オーディオ再生装置にフレーム抽出部21、評価関数演算部24、速度変換処理部23、スケールファクタ修正部25を加えたものである。フレーム抽出部21は、図18には、明示されていないが、MP3オーディオ復号化部11に含まれているものを同図では、明示的に示したものである。

【0048】フレーム抽出部21では、MP3オーディオデータのオーディオフレームとも呼ばれるフレームを抽出する機能を持ち、フレームデータをスケールファクタ抽出部22と速度変換処理部23に出力する。スケールファクタ抽出部22では、フレームからスケールファクタを抽出し、評価関数演算部24に出力する。速度変換処理部24では、フレームの間引きや繰り返しを行うと共に、評価関数によって無音区間のデータ量を削除し、スケールファクタ修正部25に出力する。スケール

ファクタ修正部25では速度変換部23で接続されたフレームの前後のスケールファクタを修正し、MPEGオーディオ復号化部26に出力する。

【0049】本構成例においては、図18に示す一般的なMPEGオーディオ再生装置に話速変換系回路22、23、24、25を追加するのみであり、話速変換機能を容易に備えることができる利点がある。

【0050】図10は、本発明の話速変換を適用したMPEGデータ再生装置の別の構成例を示す図である。図10の構成においては、図18に示める一般的なMPEGオーディオ再生装置に評価関数演算部33、話速変換処理部34、スケールファクタ修正部35を加えたものである。MPEGオーディオ復号化部31では、既にフレーム抽出機能とスケールファクタ抽出機能を持つので、本発明の実施形態の話速変換方法に必要な処理の一部を含んでいる。よって、MPEGオーディオ復号化部31のフレーム抽出機能とスケールファクタ抽出機能の利用することで回路規模を小さくできる利点がある。

【0051】MPEGオーディオ復号化部11で抽出されたフレームとスケールファクタは、評価関数演算部33に送られ、評価関数の算出が評価関数演算部33においてなされる。評価関数値とフレームは、話速変換処理部34に送られ、フレームの間引きや繰り返しに使用される。そして、話速変換されたフレームとスケールファクタは、MPEGオーディオ復号化部11に送られる。また、MPEGオーディオ復号化部12からは、スケールファクタがスケールファクタ修正部35にも送られ、スケールファクタ修正部35において、スケールファクタの修正が行われる。修正されたスケールファクタは、MPEGオーディオ復号化部11に入力される。MPEGオーディオ復号化部11は、話速変換されたフレームと修正されたスケールファクタからオーディオフレームからなるMPEGオーディオデータをデコードし、オーディオ出力部12にデコードされたデータを送る。このようにして、話速変換された音声データがオーディオ出力部12から出力される。

【0052】図11は、本発明の別の実施形態を説明する図である。なお、図11においては、図9と同じ構成要素には同じ参照番号を付して説明を省略する。

【0053】同図は、話速変換を適用したMPEGデータ再生装置を示している。これは、従来のMPEGデータ再生装置（構成要素40、41、42、43、44、45）のMPEGオーディオ復号化部を本発明の実施形態に示したMPEGオーディオデータ再生装置のMPEGオーディオ入力部とオーディオ出力部を除いた回路と置き換えたものである。従って、上記実施形態と同様の利点を得ることができる。

【0054】図11の構成は、MPEGデータとして、音声データのみではなく、画像データも含まれる場合の装置構成を示している。まず、MPEGデータ入力部4

0から入力されたMPEGデータは、MPEGデータ分離部41によって分離され、MPEGビデオデータとMPEGオーディオデータとに分離される。MPEGビデオデータは、MPEGビデオ復号化部42に入力され、MPEGオーディオデータは、フレーム抽出部21に入力される。MPEGビデオでは、MPEGビデオ復号化部42においてデコードされ、ビデオ出力部44から出力される。

【0055】MPEGオーディオデータは、図9で説明したような処理を受け、最終的にMPEGオーディオ復号化部43において復号され、オーディオ出力部45から出力される。

【0056】図12は、本発明の更に別の実施形態である話速変換を適用したMPEGデータ再生装置の一構成例を示す図である。なお、同図において、図10、11と同じ構成要素には同じ参照番号を付し、説明を省略する。

【0057】図12の構成は、従来のMPEGデータ再生装置のMPEGオーディオ復号化部を図10に示したMPEGオーディオデータ再生装置のMPEGオーディオ入力部とオーディオ出力部を除いた回路と置き換えたものである。従って、図10の構成例と同様の利点を得ることができる。

【0058】すなわち、MPEGデータ分離部41で分離されたMPEGオーディオデータには、MPEGオーディオ復号化部43において、フレーム抽出、スケールファクタ抽出処理が施され、これらの処理結果は、評価関数演算部33とスケールファクタ修正部35に入力されて、上記したような処理により、話速変換が行われる。

【0059】図13は、本発明の更に別の実施形態であり、MPEGデータ再生装置を示す図である。なお、同図においては、図12と同じ構成要素には同じ参照番号を付してある。

【0060】図13においては、従来のMPEGデータ再生装置に評価関数演算部33、データ蓄積部50、入力データ選択部51、出力データ選択部52を加えたものである。特に、上記構成例では、MPEGオーディオデータの処理のみを独立に考えていたが、図13においては、画像データと音声データの両方を速度変換する構成を示している。

【0061】ここで、評価関数演算はMPEGオーディオ復号化部43あるいはMPEGビデオ復号化部42から種々のパラメータを得て評価関数を演算するものであり、データ蓄積部50はMPEGデータを蓄積するものであり、入力データ選択部51は評価関数と、所定の法則に従って、MPEGデータ蓄積部50から入力するMPEGデータを選択するものであり、出力データ選択部52は評価関数と所定の法則に従って出力するデータを選択するものである。

【0062】また、ユーザからの再生速度指定は、評価関数演算部33に入力され、その再生速度情報が入力データ選択部51に通知される。ここで、評価関数のパラメータとしては、例えば、再生速度、スケールファクタ、オーディオフレームカウントなど話速変換用パラメータ、また、音圧、セリフなどの音声から得られる情報、ビデオフレームカウント、フレームレート、色情報、離散コサイン変換直流成分、動きベクトル、シーンチェンジ、字幕などの映像から得られる情報が有効である。回路規模の比較的大きいフレームメモリや画像演算回路はコスト上昇につながるため、使用を避ける場合には、これらの内、ビデオフレームカウント、フレームレート、離散コサイン変換直流成分、動きベクトルなどデコードなしに得られる情報を評価関数のパラメータに用いても良い。また、MPEGビデオ復号化部42にシーンチェンジ検出機能があれば、本発明の実施形態における話速変換機能と組み合わせることで、つまり、シーンチェンジフレームとスケールファクタと再生速度から評価関数を演算することで、無音区間のシーンの欠落無しに話速変換されたダイジェスト映像を出力することが可能となる。

【0063】通常の再生時では、MPEGデータ蓄積部50からは連続的にMPEGデータが読みだされるが、例えば、再生速度が上限を超えるデータ転送レートを求めるような場合は再生が滞ってしまう。そこで、入力データ選択部51は、このような場合に評価関数を元に、あらかじめ読み出し不要なMPEGデータをスキップする。つまり不連続的に読み出しアドレスを決定する。具体的には評価関数によって再生されるべきビデオフレーム、オーディオフレームを決定し、再生されるべきMPEGデータのアドレスを算出する。MPEGデータにおいてオーディオを含むパケットあるいはビデオを含むパケットはパケットヘッダより判定する。MPEGオーディオについてはフレーム単位でアクセスが可能であり、フレームのデータ長はレイヤI、IIでは一定であるので、容易にアドレスを決定することができる。MPEGビデオについては、複数のフレームをひとまとまりとしたGOP単位でアクセスする。

【0064】ここで、MPEGデータは規格上MPEGオーディオデータはフレーム単位のアクセスが可能であるが、MPEGビデオデータはフレーム単位ではなく、複数のフレームをひとまとまりとしたGOP単位にアクセス可能である。しかしながら、評価関数によっては出力不要なフレームがある。そこで、出力データ選択部52は、このような場合に、評価関数を元に出力フレームを決定する。また、出力するビデオフレームとオーディオフレームの同期を調整する。

【0065】再生速度が高い場合、人間が映像と音声の同期を敏感に知覚にくくなるので、厳密な同期が不要であると考えられる。従って、映像と音声が大局的に同

期が合うように、映像をGOP単位、音声はオーディオフレーム（オーディオフレーム）単位で出力データを選択する。

【0066】図14は、本発明の実施形態をプログラムで実現する場合に必要なとされる装置のハードウェア構成例を示す図である。CPU61は、バス60を介して、ROM62、RAM63、通信インターフェース64、記録装置67、記録媒体読み取り装置68、入出力装置70に接続される。

10 【0067】ROM63には、BIOSなどが格納され、CPU61がこれを実行することによって、CPU61へのユーザの指示が入出力装置70から入力可能としたり、CPU61の演算結果をユーザに提示可能とする。ここで、入出力装置は、ディスプレイ、マウス、キーボードなどからなる。

【0068】また、本発明の実施形態にかかる話速変換処理を伴うMPEGデータ再生処理を実現するプログラムは、ROM62、RAM63、記録装置67、可搬記録媒体69に格納することが可能である。ROM62、RAM63に格納された当該プログラムは、CPU61が直接実行する。記録装置67や可搬記録媒体69に記録された当該プログラムは、記録装置67からは、バス60を介して直接RAM63に展開され、可搬記録媒体69からは、記録媒体読み取り装置68で可搬記録媒体69上のプログラムを読みとって、バス60を介してRAM63に展開することによって、CPU61が実行可能な状態とする。

30 【0069】ここで、記録装置67は、ハードディスクなどであり、可搬記録媒体69は、CD-ROMやフロッピー（登録商標）ディスク、DVDなどである。また、装置は、通信インターフェース64を備えてもよく、この場合、ネットワーク65を介して情報提供者66のデータベースにアクセスし、当該プログラムをダウンロードして使用することの可能である。あるいは、ネットワーク65がLANのように構築されている場合には、ネットワーク環境下で当該プログラムを実行することも可能である。

<付記>

1. オーディオデータを含む圧縮されたマルチメディアデータを再生する装置であって、該オーディオデータの単位データであるフレームを抽出する抽出手段と、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行う変換手段と、該変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備えた、ことを特徴とするデータ再生装置。

2. オーディオデータを含むマルチメディアデータを再生する装置であって、該オーディオデータの単位データであるフレームを抽出する抽出手段と、該オーディオデータの再生速度を設定する設定手段と、該設定手段によ

って設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行う速度変換手段と、該速度変換手段から受け取った該オーディオデータの該フレームをデコードし、音声を再生する再生手段とを備え、圧縮されたオーディオデータをデコードすることなしに再生速度変換処理を行うことを特徴とするデータ再生装置。

3. 前記オーディオデータは、MPEGオーディオデータであることを特徴とする付記2に記載のデータ再生装置。

4. 前記フレームに含まれるスケールファクタを抽出するスケールファクタ抽出手段と、該スケールファクタに対して演算を行う演算手段と、該演算手段の演算結果と所定の閾値とを比較し、該演算結果が該閾値よりも小さい場合には、対応する該フレームを前記再生手段に送らないように制御する制御手段と、を更に備えることを特徴とする付記3に記載のデータ再生装置。

5. 前記演算手段は、前記フレームに含まれる複数のスケールファクタの和を演算することを特徴とする付記4に記載のデータ再生装置。

6. 前記フレームに含まれる複数のスケールファクタの内、前記再生手段によって再生されるときに、フレーム間の接続部分に生じる音圧の不連続な変化をうち消すためのスケールファクタ修正係数を生成し、該スケールファクタと該スケールファクタ修正係数を演算して、該再生手段にデコードすべきデータとして入力するスケールファクタ修正手段を更に備えることを特徴とする付記4に記載のデータ再生装置。

7. 前記データ再生装置は、画像データとオーディオデータを含むマルチメディアデータを受信し、更に、画像データとオーディオデータを分離する分離手段と、該画像データを復号する復号手段と、該画像データを再生する画像再生手段とを備えることを特徴とする付記2に記載のデータ再生装置。

8. 前記画像データとオーディオデータは、MPEGデータとして構成されていることを特徴とする付記7に記載のデータ再生装置。

9. オーディオデータを含むマルチメディアデータを再生する方法であって、

(a) 該オーディオデータの単位データであるフレームを抽出するステップと、

(b) 該オーディオデータの再生速度を設定するステップと、

(c) 該ステップ(b)において設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行うステップと、

(d) 該ステップ(c)の処理後受け取った該オーディオデータの該フレームをデコードし、音声を再生するステップとを備え、圧縮されたオーディオデータをデコー

ドすることなしに再生速度変換処理を行うことを特徴とするデータ再生方法。

10. 前記オーディオデータは、MPEGオーディオデータであることを特徴とする付記9に記載のデータ再生方法。

11. (e) 前記フレームに含まれるスケールファクタを抽出するステップと、

(f) 該スケールファクタに対して演算を行うステップと、

10 (g) 該ステップ(f)の演算結果と所定の閾値とを比較し、該演算結果が該閾値よりも小さい場合には、対応する該フレームに前記ステップ(d)の処理を行わないように制御するステップと、を更に備えることを特徴とする付記10に記載のデータ再生方法。

12. 前記ステップ(f)は、前記フレームに含まれる複数のスケールファクタの和を演算することを特徴とする付記11に記載のデータ再生方法。

13. (h) 前記フレームに含まれる複数のスケールファクタの内、前記ステップ(d)において再生されるときに、フレーム間の接続部分に生じる音圧の不連続な変化をうち消すためのスケールファクタ修正係数を生成し、該スケールファクタと該スケールファクタ修正係数を演算した値に基づいて、該ステップ(d)の処理を行うステップを更に備えることを特徴とする付記11に記載のデータ再生方法。

14. 前記データ再生方法は、画像データとオーディオデータを含むマルチメディアデータを処理し、

更に、

(i) 画像データとオーディオデータを分離するステップと、

(j) 該画像データを復号するステップと、

(k) 該画像データを再生するステップとを備えることを特徴とする付記9に記載のデータ再生方法。

15. 前記画像データとオーディオデータは、MPEGデータとして構成されていることを特徴とする付記14に記載のデータ再生方法。

16. 情報装置に、オーディオデータを含むマルチメディアデータを再生する方法を実行させるプログラムを記録した記録媒体であって、

40 (a) 該オーディオデータの単位データであるフレームを抽出するステップと、

(b) 該オーディオデータの再生速度を設定するステップと、

(c) 該ステップ(b)において設定された該再生速度に基づいて、該オーディオデータの該フレームの間引き処理あるいは該フレームの繰り返し出力処理を行うステップと、

50 (d) 該ステップ(c)の処理後受け取った該オーディオデータの該フレームをデコードし、音声を再生するステップとを備え、圧縮されたオーディオデータをデコー

ドすることなしに再生速度変換処理を行うことを特徴とする方法を実現するプログラムを記録した、情報装置読み取り可能な記録媒体。

17. 前記オーディオデータは、MPEGオーディオデータであることを特徴とする付記16に記載の記録媒体。

18. (e) 前記フレームに含まれるスケールファクタを抽出するステップと、

(f) 該スケールファクタに対して演算を行うステップと、

(g) 該ステップ(f)の演算結果と所定の閾値とを比較し、該演算結果が該閾値よりも小さい場合には、対応する該フレームに前記ステップ(d)の処理を行わないように制御するステップと、を更に備えることを特徴とする付記17に記載の記録媒体。

19. 前記ステップ(f)は、前記フレームに含まれる複数のスケールファクタの和を演算することを特徴とする付記18に記載の記録媒体。

20. (h) 前記フレームに含まれる複数のスケールファクタの内、前記ステップ(d)において再生されるときに、フレーム間の接続部分に生じる音圧の不連続な変化をうち消すためのスケールファクタ修正係数を生成し、該スケールファクタと該スケールファクタ修正係数を演算した値に基づいて、該ステップ(d)の処理を行うステップを更に備えることを特徴とする付記18に記載の記録媒体。

21. 前記データ再生方法は、画像データとオーディオデータを含むマルチメディアデータを処理し、更に、

(i) 画像データとオーディオデータを分離するステップと、

(j) 該画像データを復号するステップと、

(k) 該画像データを再生するステップとを備えることを特徴とする付記16に記載の記録媒体。

22. 前記画像データとオーディオデータは、MPEGデータとして構成されていることを特徴とする付記21に記載の記録媒体。

【0070】

【発明の効果】以上説明したように、本発明によれば、MPEGオーディオ規格で定義されるフレーム単位に処理することでMPEGデータをデコードせずに話速を変化させる話速変換する効果を奏し、スケールファクタを用いることでMPEGデータをデコードせずに無音区間を圧縮して話速変換する効果を有する。

【0071】また、フレーム間の接続部の前後のスケールファクタを変化させることで、フレームの接続部での聴感上の違和感を低減する効果を奏し、MPEGデータ再生方法及びMPEGデータ再生装置の性能向上に寄与するところが大い。

【図面の簡単な説明】

【図1】同じ音声データをMPEGオーディオレイヤIで圧縮したデータのスケールファクタと圧縮しないデータの音圧について比較した図である。

【図2】本発明の話速変換処理の概略フローチャートである。

【図3】本発明の別の話速変換処理の概略フローチャートである。

【図4】再生速度変換処理の流れを示す詳細なフローチャートである。

10 【図5】再生速度変換処理及び無音部分除去処理を含む処理の流れを示す詳細フローチャートである。

【図6】ノイズ低減処理の流れを示すフローチャートである。

【図7】図6のスケールファクタ修正処理を説明する図(その1)である。

【図8】図6のスケールファクタ修正処理を説明する図(その2)である。

【図9】本発明の話速変換を適用したMPEGオーディオデータ再生装置の一構成を示すブロック構成図である。

【図10】本発明の話速変換を適用したMPEGデータ再生装置の別の構成例を示す図である。

【図11】本発明の別の実施形態を説明する図である。

【図12】本発明の更に別の実施形態である話速変換を適用したMPEGデータ再生装置の一構成例を示す図である。

【図13】本発明の更に別の実施形態であり、MPEGデータ再生装置を示す図である。

【図14】本発明の実施形態をプログラムで実現する場合に必要とされる装置のハードウェア構成例を示す図である。

【図15】MPEGオーディオデータのフォーマットを示す図(その1)である。

【図16】MPEGオーディオデータのフォーマットを示す図(その2)である。

【図17】従来のMPEGオーディオ再生装置の基本的構成を示す図である。

【符号の説明】

10、20 MPEGオーディオ入力部

11、26、43 MPEGオーディオ復号化部

12、27、45 オーディオ出力部

21 フレーム抽出部

22 スケールファクタ抽出部

23、34 速度変換部(話速変換処理部)

24、33 評価関数演算部

25、35 スケールファクタ修正部

40 MPEGデータ入力部

41 MPEGデータ分離部

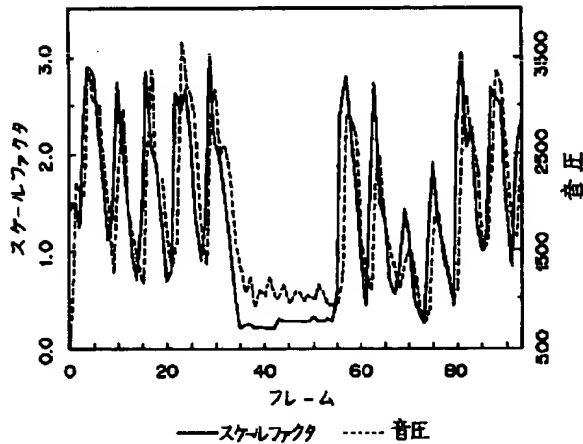
42 MPEGビデオ復号化部

50 44 ビデオ出力部

50 MPEGデータ蓄積部
51 入力データ選択部

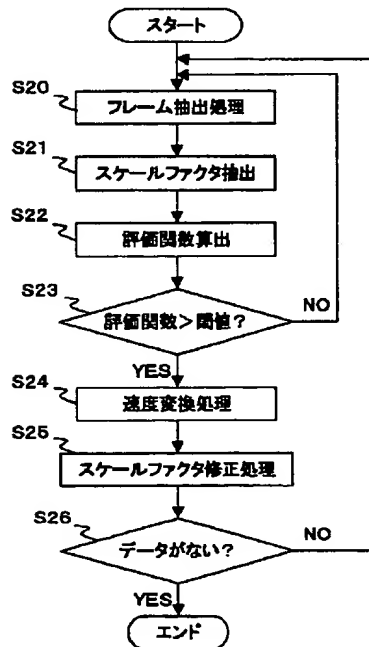
【図1】

同じ音声データを MPEGオーディオレイヤIIで圧縮した
データのスケールファクタと圧縮しないデータの音圧について
比較した図



【図3】

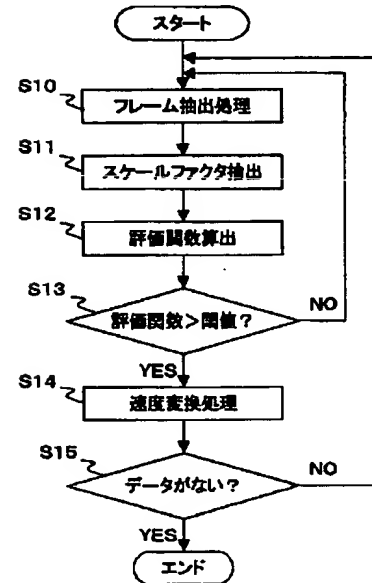
本発明の別の話速変換処理の概略フローチャート



52 出力データ選択部

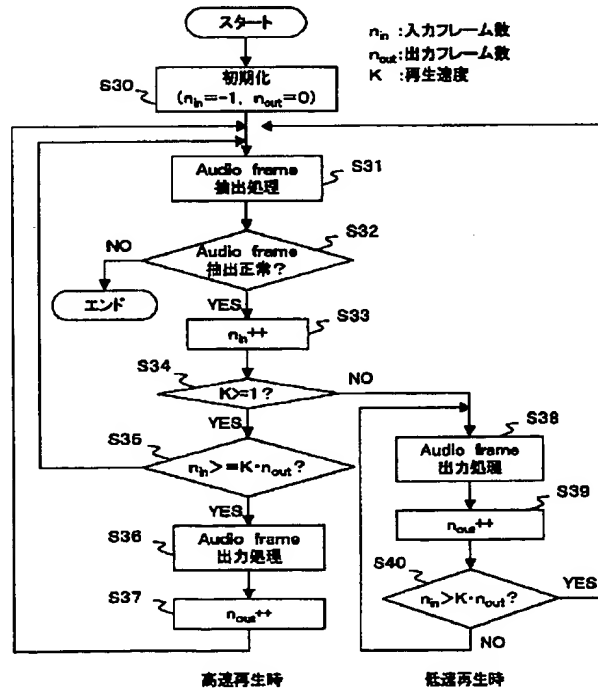
【図2】

本発明の話速変換処理の概略フローチャート



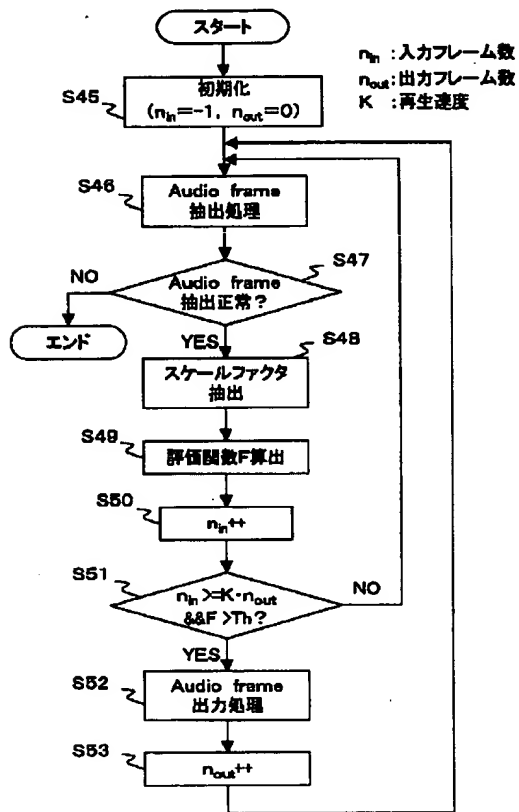
【図4】

再生速度変換処理の流れを示す詳細なフローチャート



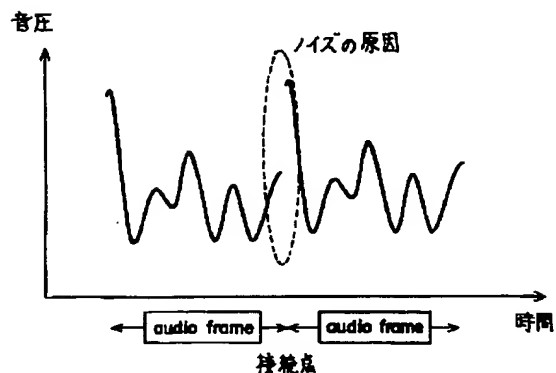
【図5】

再生速度変換処理及び無音部分除去処理を含む
処理の流れを示す詳細フローチャート



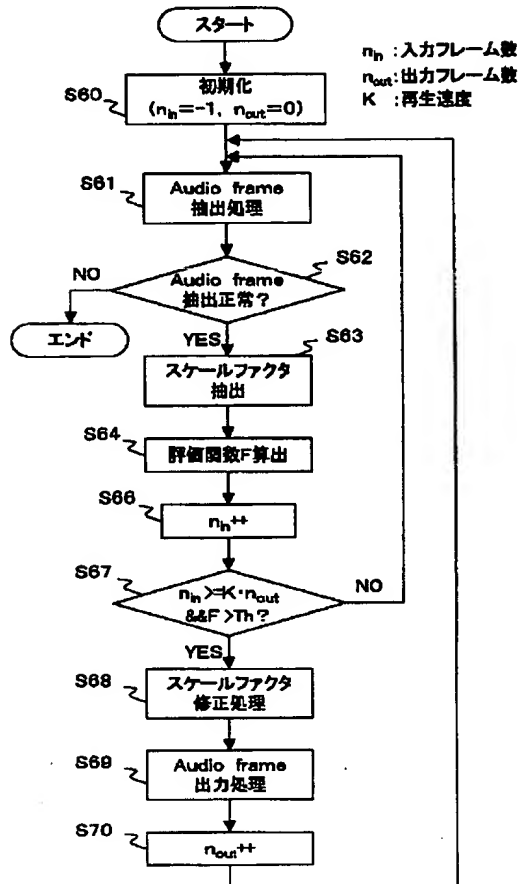
【図7】

図6のスケールファクタ修正処理を説明する図
(その1)



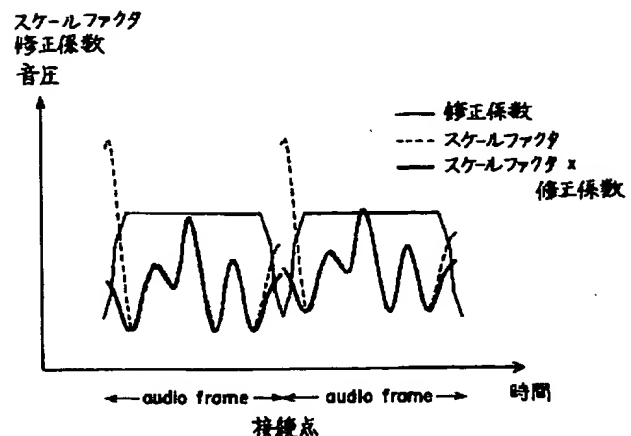
【図6】

ノイズ低減処理の流れを示すフローチャート



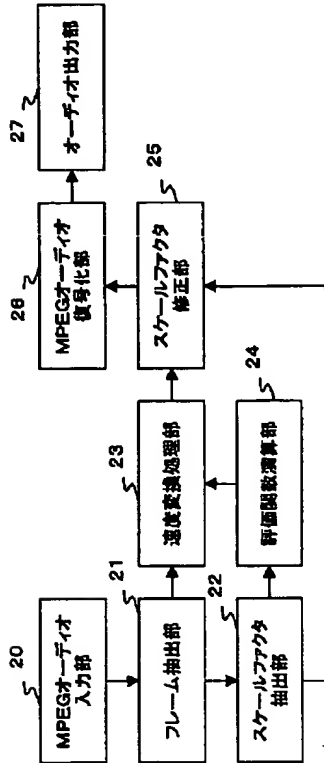
【図8】

図6のスケールファクタ修正処理を説明する図
(その2)



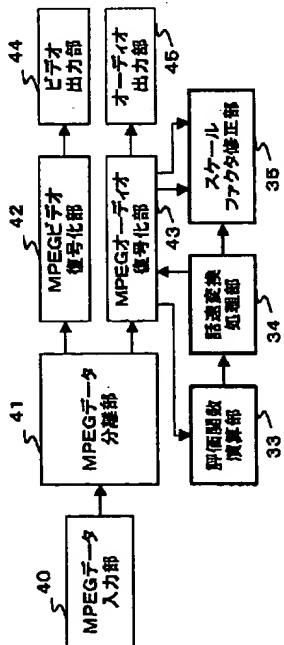
【図9】

本発明の低速変換を適用したMPEGオーディオデータ再生装置の構成例を示すブロック構成図



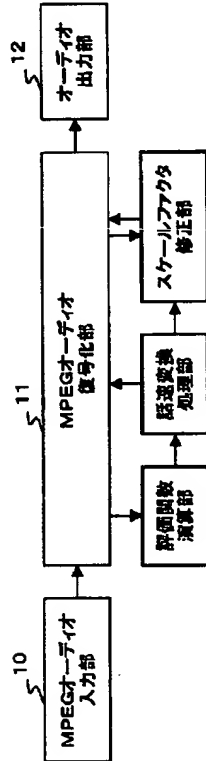
【図12】

本発明の更に別の実施形態である
低速変換を適用したMPEGデータ
再生装置の構成例を示す図



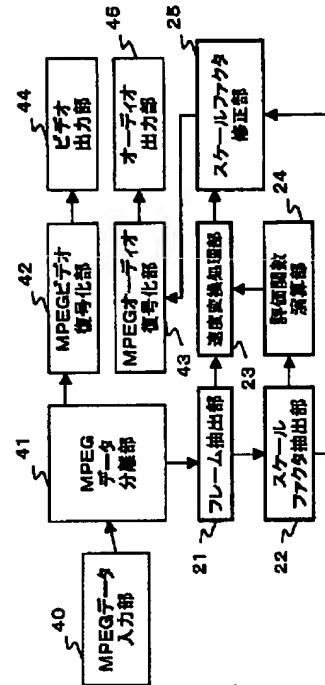
【図10】

本発明の低速変換を適用したMPEGデータ
再生装置の別の構成例を示す図



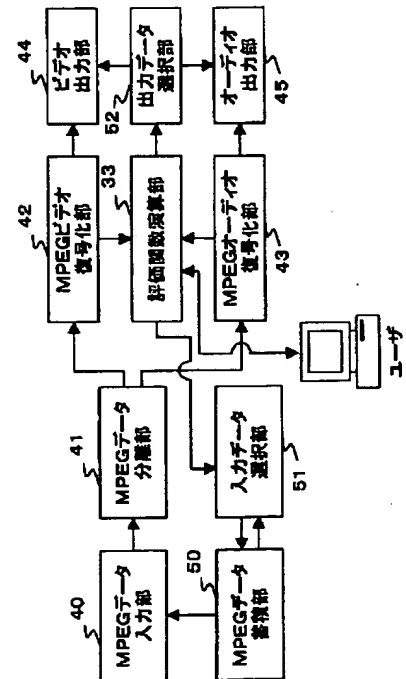
【図11】

本発明の別の実施形態を説明する図



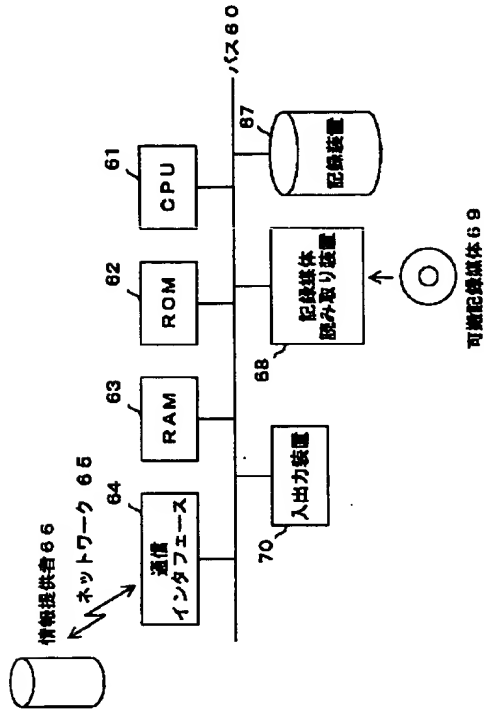
【図13】

本発明の更に別の実施形態であり、
MPEGデータ再生装置を示す図



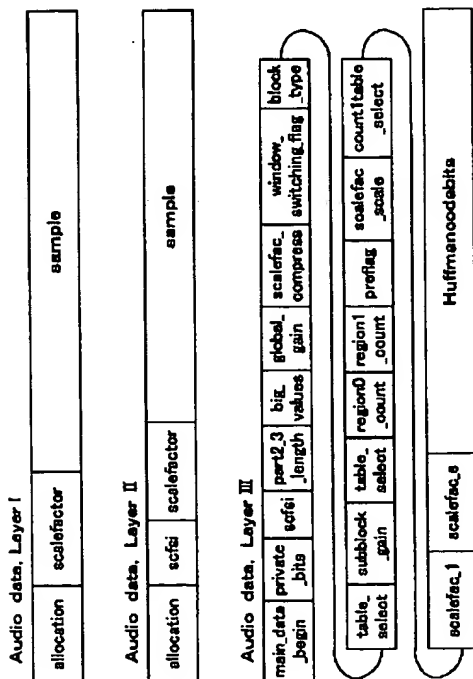
【図14】

本発明の実施形態をプログラムで実現する場合に必要なとされる装置のハードウェア構成例を示す図



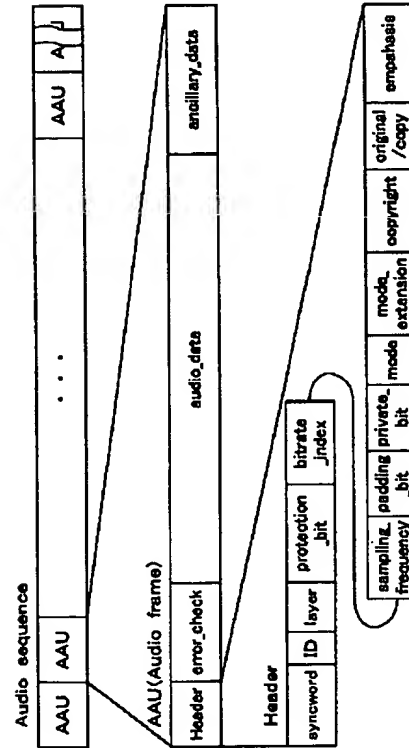
【図16】

MPEGオーディオデータのフォーマットを示す図 (その2)



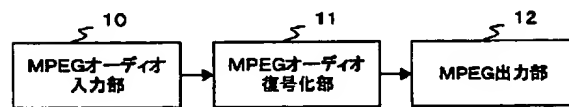
【図15】

MPEGオーディオデータのフォーマットを示す図 (その1)



【図17】

従来のMPEGオーディオ再生装置の基本的構成を示す図



フロントページの続き

(72)発明者 肥塚 哲雄
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

Fターム(参考) 5D044 AB05 DE03 DE14 FG23 GK02
GK08
5D045 DA00